UNIVERSITY of York

This is a repository copy of *Structural matching by discrete relaxation*.

White Rose Research Online URL for this paper:
https://eprints.whiterose.ac.uk/1985/

Article:
Wilson, R C orcid.org/0000-0001-7265-3033 and Hancock, E R orcid.org/0000-0003-4496-2028 (1997) Structural matching by discrete relaxation. IEEE Transactions on Pattern Analysis and Machine Intelligence. pp. 634-648. ISSN 0162-8828

https://doi.org/10.1109/34.601251

Reuse
Items deposited in White Rose Research Online are protected by copyright, with all rights reserved unless indicated otherwise. They may be downloaded and/or printed for private study, or other acts as permitted by national copyright laws. The publisher or other rights holders may allow further reproduction and re-use of the full text version. This is indicated by the licence information on the White Rose Research Online record for the item.

Takedown
If you consider content in White Rose Research Online to be in breach of UK law, please notify us by emailing eprints@whiterose.ac.uk including the URL of the record and the reason for the withdrawal request.

White Rose university consortium
Universities of Leeds, Sheffield & York

eprints@whiterose.ac.uk
https://eprints.whiterose.ac.uk/

Structural Matching by Discrete Relaxation

Richard C. Wilson and Edwin R. Hancock

Abstract—This paper describes a Bayesian framework for performing relational graph matching by discrete relaxation. Our basic aim is to draw on this framework to provide a comparative evaluation of a number of contrasting approaches to relational matching. Broadly speaking there are two main aspects to this study. Firstly we focus on the issue of how relational inexactness may be quantified. We illustrate that several popular relational distance measures can be recovered as specific limiting cases of the Bayesian consistency measure. The second aspect of our comparison concerns the way in which structural inexactness is controlled. We investigate three different realizations of the matching process which draw on contrasting control models. The main conclusion of our study is that the active process of graph-editing outperforms the alternatives in terms of its ability to effectively control a large population of contaminating clutter.

Index Terms—Structural graph matching, discrete relaxation, energy minimization, Bayesian, graph edit, clutter, MAP estimation, SAR images, infrared images.

1 Introduction

Relational structure matching has been a task of pivotal importance since the inception of machine vision. It was the seminal paper of Barrow and Popplestone [1] that first established the relational graph as a practical representation for scene matching. Moreover, the subsequent paper of Barrow and Burstall [2] presented some viable structural matching techniques based upon the idea of searching for maximal cliques in the association graph. More recently, Horaud and Skordas [20] have exploited these concepts in the domain of stereo correspondence. However, these early structural matching techniques are only effective when exact relational descriptions are to hand. Effective relational matching must be capable of accommodating two classes of error. The first of these are initialization errors which result from ambiguities in object appearance. These ambiguities may be the consequence of either unary attribute variations or uncertainties in the measurement acquisition process. The second source of error is structural disturbance of the relational descriptions under match. These structural errors result from either poor initial image segmentation or the presence of noise and clutter. Based on these observations, there are two important modeling issues. The first of these is the way in which consistency of match is measured. Perhaps of even greater importance for practical vision systems, is the way in which structural corruption caused by clutter or noise is identified. In other words, the goal is not only the location of optimally consistent matches, it is also the rectification of structural errors.

Because of the pivotal importance of relational matching in high and intermediate level vision, a diverse family of algorithms have been developed with the aim of meeting these dual goals of gauging consistency and overcoming structural errors. For instance, the quest for a measure of relational consistency has been central to the work of Shapiro and Haralick [27] and of Boyer and Kak [3]. The control of structural errors was not only the motivation behind Barrow and Popplestone's [1] idea of searching for maximal subgraphs of the association graph, it also led Sanfeliu and Fu [24] to their graph edit process. However, although individually effective, the available algorithms originate from distinct methodological foundations. In consequence their relationship to one-another is not easily appreciated. Moreover, there has been little attempt to provide any objective experimental comparison. Our aim in this paper is therefore to provide both a theoretical and experimental comparison of a number of contrasting matching techniques. The framework for this study is a recently reported Bayesian approach to relational matching that is capable of both gauging consistency and rectifying structural errors [31], [32].

1.1 Related Literature

It was precisely the observation that effective relational matching is limited by both initialization error and structural error that stimulated much of the classical work on structural pattern recognition of the 1980s. In a nutshell, this work aimed to realize relational matching by inexact means. Shapiro and Haralick not only identified the difficulties associated with the search for subgraph isomorphisms when the graphs under match are corrupted realizations of one-another [26], but also proposed a relational distance metric which could be used to gauge structural differences [27]. According to Shapiro and Haralick [26], missing or extraneous entities are accommodated by inserting null nodes into the graphs without penalty. A similar idea was pursued by Sanfeliu and Fu [24] who propose a distance measure based on the number of structural edit operations needed to transform corrupted graphs into their exact equivalents. The edit operations include node relabeling, node deletion or insertion together with an analo-

• *The authors are with Department of Computer Science, University of York, York, Y01 5DD, England. E-mail: {wilson; erh}@minster.york.ac.uk.*

Manuscript received 7 Mar. 1996; revised 21 Apr. 1997. Recommended for acceptance by L. Shapiro.
For information on obtaining reprints of this article, please send e-mail to: transpami@computer.org, and reference IEEECS Log Number 104924.

0162-8828/97/$10.00 © 1997 IEEE

gous set of operations for edges. Associated with each class of operation is an associated cost. Structural matching attempts to minimize the global cost. This graph edit idea has recently been extended by Messmer and Bunke [22] who develop an error tolerant technique for locating subgraph isomorphisms in a large model base. Wong and You [34] have taken some steps towards placing the inexact graph matching problem on an information theoretic footing by introducing the concept of the entropy associated with a structural match. Some of the practical advantages of structural matching have been demonstrated by Kim and Kak [21] who show how discrete relaxation can be used to match bi-partite graphs representing range-images and CAD models.

Despite the progress made in the structural matching of inexact graphs, one of the perceived limitations of the approach has been its exclusively symbolic nature and its failure to draw on the wealth of measurement information available in realistic matching problems. This shortcoming has been addressed by Boyer and Kak [3] who develop a structural matching technique that measures attribute deviations between graphs using the logarithmic conditional information. Christmas, Kittler, and Petrou [5] draw on a Gaussian model of binary attribute relations to compute compatibility coefficients in a Bayesian relaxation scheme. This evidence combining approach is a probabilistic relaxation scheme which represents a significant enhancement of the ideas originally pioneered by Rosenfeld, Hummel, and Zucker [23] and first exploited in the graph-matching domain by Faugeras and Price [7]. Yang and Kittler [35] extend these ideas to higher order attribute relations using the apparatus of mean-field theory [13]. Notwithstanding their successes, these attribute-based methods place heavy demands upon the reliable estimation of the underlying model parameters [4], [8], [33].

Interest in the matching of symbolic structures has recently received a renewed impetus with the success of aspect graphs as a hierarchical representation of 3D shape [6], [10] and with the need to index into large model-bases [25]. Moreover, in a recent series of papers [31], [32] we have demonstrated that the rejection of structural methods as an effective means of matching may be regarded as premature. Basic to our recent work is the idea of developing a finer way of gauging the cost of symbolic consistency using an explicit model of the various types of noise and segmentation error anticipated to be present. Key to the approach is the construction of a Bayesian model of relational inexactness. The underlying model is extremely simple and assumes that matching errors occur with a uniform and memoryless probability distribution. This model leads to an exponential measure of relational consistency which draws on the Hamming distance between the structural units of the graphs under match.

1.2 Paper Overview

The techniques described above represent a disparate family of algorithms for measuring relational consistency and correcting structural errors. However, they originate from distinct methodological foundations, and in consequence their relationship to one-another is not easily appreciated. Moreover, there has been little attempt to provide any objective experimental algorithm comparison. Our aim in this paper is to take advantage of our Bayesian framework [31], [32] to provide both a theoretical and experimental comparison.

We commence our study by showing how the exponential Bayesian consistency measure approximates the conventional relational distance metric [27] and the action of a Hopfield associative memory [18], [28], [29] under suitable limiting conditions. The conventional relational distance measure corresponds to conditions when the uniform probability of matching errors is vanishingly small; the Hopfield memory corresponds to the case when the match error-probability is close to 1/2. Although each of these approximate methods is inferior in performance to the compound exponential functional, we are able to establish operating conditions under which they prove to be most effective.

Turning our attention to the control of clutter, there are a number of ways in which this new measure of relational consistency may be exploited to enhance the matching process when noise is a limiting factor. The most straightforward of these is to attempt to locate the most consistent matches while simultaneously labeling clutter in an optimization process. This is the strategy adopted by Boyer and Kak [3], and, by Christmas, Kittler, and Petrou [4]. Here the structure of the graphs remains unmodified in the matching process; in other words, extraneous nodes are simply tagged with a null-label. A second strategy is to use the consistency measure to identify candidate matches for inclusion in an association graph; clutter may then be rejected by employing a constraint filtering process [2], [31] similar to that originally described by Barrow and Popplestone [1]. This is a hybrid technique which combines optimization ideas with classical constraint filtering techniques. A third, and novel alternative, is to use the consistency measure to control the reconfiguration of the graphs in an active matching strategy [32] using graph-edit operations. In this way clutter is excluded by deleting nodes and reconfiguring the edge set of the graphs. Although this process has many of the conceptual ingredients of the graph edit process of Sanfeliu and Fu [24], it differs in two important respects. Firstly, the Bayesian measure of relational consistency implicitly gauges the cost of structural errors using only a single model parameter. Secondly, the MAP framework directly couples the structural edit operations to the attributes of the raw image entities. Our second aim in this paper is therefore to show how these alternative strategies for controlling clutter can be accommodated within our framework for modeling relational consistency. Moreover, we offer some comparison of their relative robustness to noise and clutter.

The outline of this paper is as follows. Section 2 introduces the formal ingredients of our method and describes our abstraction of scenes in terms of relational graphs. The MAP (maximum a posteriori probability) estimation framework that underpins our matching technique is reviewed in Section 3. In Section 4 we describe our approach to modeling relational consistency using the concept of a label error process. Section 5 develops various approxima-

tions of the exponential consistency measure and identifies the relationship to various alternatives described elsewhere in the literature. Section 6 explores different strategies for controlling clutter and noise. Section 7 presents a comparative sensitivity analysis of both the different consistency measures and the different clutter control mechanisms. Real-world experiments are presented in Section 8. Finally, Section 9 presents some conclusions and suggests directions for further investigation.

2 Relational Graphs

We abstract the matching process in terms of attributed relational graphs [3], [10], [4], [27]. We use the triple $G = (V, E, \mathcal{A})$ to denote the graphs under match, where V is the set of nodes, E is the set of arcs, and $\mathcal{A} = \{\underline{x}_i, \forall i \in V\}$ is a set of unary measurements associated with the nodes. Our aim in matching is to associate nodes in a graph $G_1 = (V_1, E_1, \mathcal{A}_1)$, where $\mathcal{A}_1 = \left\{\underline{x}_u^1, \forall u \in V_1\right\}$, representing data to be matched against those in a graph $G_2 = (V_2, E_2, \mathcal{A}_2)$, where $\mathcal{A}_2 = \left\{\underline{x}_v^2, \forall v \in V_2\right\}$, representing an available relational model. This matching process is facilitated using constraints provided by suitable relational subunits of the model graph G_2. Unmatchable nodes originating from noise or clutter elements are accommodated by augmenting the model-graph nodes with a null-label ϕ. In other words, this null-label may be used to tag data-graph nodes for which no acceptable match exists in the model. Formally, the matching is represented by a function $f: V_1 \rightarrow V_2 \cup \phi$ from the nodes in the data graph G_1 to those in the augmented model graph G_2. The function f consists of a set of Cartesian pairs drawn from the space of possible matches between the two graphs, i.e., $f \subseteq V_1 \times (V_2 \cup \phi)$. The function provides a convenient device for indexing the nodes in the data graph G_1 against their matched counterparts in the model graph G_2. We use the notation $(u, v) \in f$ to denote the match of node $u \in V_1$ against node $v \in V_2 \cup \phi$.

In performing the matches of the nodes in the data graph G_1 we will be interested in exploiting structural constraints provided by the model graph G_2. We use subgraphs that consist of neighborhoods of nodes interconnected by arcs. For convenience we refer to these structural subunits or N-ary relations as super-cliques. The super-clique of the node indexed j in the graph G_1 with arc-set E_1 is denoted by the set of nodes $C_j = j \cup \{i \mid (i, j) \in E_1\}$. We use the notation

$$R_j = \left(u_1, u_2, \ldots, u_{|C_j|}\right)$$

to denote the N-ary symbolic relation represented by the nodes of the super-clique $C_j \subset V_1$ in the data graph G_1. The matched realization of this super-clique is denoted by the relation

$$\Gamma_j = \left(f(u_1), f(u_2), \ldots, f(u_{|C_j|})\right).$$

Our aim is to modify the match to optimize a measure of global consistency with the constraints provided by the model graph G_2. The constraints available to us are provided by the N-ary symbol relations on the super-cliques of the model graph G_2. The critical ingredient in developing our matching scheme is the set of feasible structure preserving mappings between each super-clique of graph G_1 and those of graph G_2. The set of feasible mappings, or dictionary, for the super-clique C_j is denoted by $\Theta_j = \{S_i\}$ where $S_i = i \cup \{j \mid (i, j) \in E_2\}$. Each element S_i of Θ_j is therefore a relation formed on the nodes of the model graph; we denote such consistent relations by $S_i = (v_1, v_2, \ldots)$. The dictionary of feasible mappings for the super-clique C_j consists of all the consistent relations that may be elicited from the graph G_2.

In practice, the dictionary of relations must be constructed so as to allow for super-clique size differences due to node loss and merging. We accommodate this effect by padding model graph super-cliques with sufficient dummy nodes so as to raise them to the same cardinality as their data-graph counterparts. The nodes of the padded model-graph relations are then permuted so as to preserve the cyclicity of the non-dummy nodes. An illustrative example is shown in Fig. 1. This process effectively preserves the adjacency structure of the surviving model graph nodes while leaving dictionary invariant to potential scene translations, scalings or rotations. The idea of padding relations in this way was originally suggested by Shapiro and Haralick [27], [26]. However, whereas Shapiro and Haralick's dummy nodes are inserted without penalty, our consistency model described in Section 4 associates a probability with insertions.

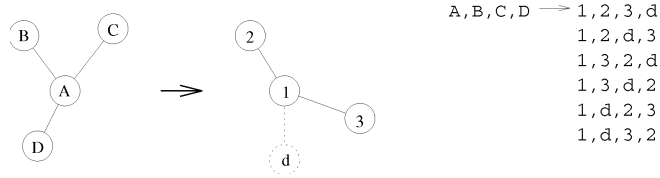

Fig. 1. Example of super-clique mapping.

3 MAP Estimation

Our aim in performing graph matching is to iteratively relabel the nodes of the data graph so as to optimize a global MAP criterion. From the standpoint of information theory, we therefore seek the matched configuration of nodes that has maximum a posteriori probability with respect to the available unary measurement information. More formally, the aim is to realize an iterative label reassignment process that optimizes the quantity

$$P(f \mid \mathcal{A}_1, \mathcal{A}_2) = \frac{p(\mathcal{A}_1, \mathcal{A}_2 \mid f) P(f)}{p(\mathcal{A}_1, \mathcal{A}_2)} \quad (1)$$

where $P(f)$ is the joint prior for the current matching configuration. The quantities $p(\mathcal{A}_1, \mathcal{A}_2 \mid f)$ and $p(\mathcal{A}_2, \mathcal{A}_2)$ are,

respectively, the conditional measurement density and the probability density function for the sets of unary measurements. We commence our development by making the assumption that the different pairs of unary measurements are conditionally independent of one-another given the current state of match. Under these conditions, the joint conditional density appearing on the right-hand side in (1) may be factorized over the set of Cartesian pairs that constitute the matching function. Suppose that $\underline{x}_u^1$ is the unary attribute residing at the node u of the data graph G_1 and that $\underline{x}_v^2$ is the unary attribute on the node v of the model graph G_2. Measurement independence allows us to factorize the joint-conditional density in the following manner over the Cartesian pairs (u, v) belonging to the matching function f

$$p(\mathcal{A}_1, \mathcal{A}_2 \mid f) = \prod_{(u,v)\in f} p\left(\underline{x}_u^1, \underline{x}_v^2 \mid u, v\right) \tag{2}$$

As a result of this assumption and applying the Bayes theorem, the above expression can be recast in terms of the a posteriori matching probabilities $P\left(u, v \mid \underline{x}_u^1, \underline{x}_v^2\right)$ and the matching priors $P(u, v)$ for the match $f(u) = v$

$$p(\mathcal{A}_1, \mathcal{A}_2 \mid f) = \prod_{(u,v)\in f} p\left(u, v \mid \underline{x}_u^1, \underline{x}_v^2\right) \frac{p\left(\underline{x}_u^1, \underline{x}_v^2\right)}{P(u, v)} \tag{3}$$

Because the unconditional measurement densities $p(\mathcal{A}_1, \mathcal{A}_2)$ and $p\left(\underline{x}_u^1, \underline{x}_v^2\right)$ are independent of the match labels and are hence a static property of the data under consideration, the local optimum of the a posteriori probability may be located by applying the following iterative decision rule to update the matching configuration

$$f(u) = \arg\max_{v\in V_2\cup\phi} \frac{P\left(u, v \mid \underline{x}_u^1, \underline{x}_v^2\right)}{P(u, v)} P(f) \tag{4}$$

In the above formula, $\arg\max_{v\in V_2\cup\phi}$ indicates that the assigned label v at the node u in the data graph G_1 is such as to maximize the a posteriori probability over the space of feasible matches $V_2 \cup \phi$.

Essentially, we are concerned with developing an iterative reassignment process for locating an optimal match that satisfies certain constraints upon relational consistency. In order to apply this global optimization strategy we require three additional ingredients. These are:

1) a model of the a posteriori matching probabilities,
2) a model of the joint prior, and
3) an effective algorithm control strategy.

The most critical of these is a means of modeling the joint priors $P(f)$ so as to gauge the overall consistency of the surviving relational structure. It is the modeling of relational consistency that is the focus of our attention in the next section of this paper. The issues of control are discussed in Section 6. Details of the modeling of a posteriori matching probabilities are application dependent and are hence deferred until we address the issue of experimental validation of the technique in Section 8.

4 Relational Consistency

In this section we describe the development of our relational consistency measure [31], [32]. The underlying modeling of structural consistency is Bayesian and commences from the joint prior for the individual super-clique matches [31], [32], i.e., $P(\Gamma_j)$. This model of the matching probability can be viewed as providing a means of imposing constraints on consistent relational matches. Rather than imposing attribute constraints, we draw on an objective Bayesian model of relational corruption which is posed at the symbolic level. This results in a consistency metric which is a compound exponential function of relational distance (Hamming distance). Underlying this metric is a purely symbolic model of relational consistency, namely, the dictionary of structure preserving mappings Θ_j. Attribute relations are the modeling province of the a posteriori matching probabilities $P\left(u, v \mid \underline{x}_u^1, \underline{x}_v^2\right)$.

As we noted in Section 2, the consistent labelings available for gauging the quality of match are represented by the set of relational mappings from C_j onto G_2, i.e., Θ_j. As demanded by the Bayes rule, we compute the probability of the required super-clique matching by expanding over the basis configurations belonging to the dictionary Θ_j

$$P\left(\Gamma_j\right) = \sum_{S_i\in\Theta_j} P\left(\Gamma_j \mid S_i\right).P\left(S_i\right) \tag{5}$$

The development of a useful relational consistency measure from this expression requires models of the processes at play in matching and of their roles in producing errors. These models are represented in terms of the joint conditional matching probabilities $P(\Gamma_j \mid S_i)$ and of the joint priors $P(S_i)$ for the consistent relations in the dictionary. In developing the required models we will limit our assumptions to the case of matching errors which are memoryless and occur with uniform probability distribution.

To commence our modeling of the conditional probabilities, we assume that the various types of matching error for nodes belonging to the same super-clique are memoryless. In direct consequence of this assumption, we may factorize the required probability distribution over the symbolic constituents of the relational mapping under consideration. As a result the conditional probabilities $P(\Gamma_j \mid S_i)$ may be expressed in terms of a product over label confusion probabilities

$$P\left(\Gamma_j \mid S_i\right) = \prod_{k=1}^{|S_i|} P\left(f\left(u_k\right) \mid v_k\right) \tag{6}$$

Our next step is to propose a two component model of the processes which give rise to erroneous matches. The first of these processes is initialization error, which we aim to rectify by iterative label updates. We assume that initialization errors occur with a uniform and memoryless probability P_e. The second source of error is structural disturbance of the relational graphs caused by noise, clutter or segmentation error. It is the rectification of these errors which poses the greatest challenge to relational matching. It

is not only the motivation of the structural edit operations of Sanfeliu and Fu [24], but also of the null match process of both Boyer and Kak [3] and Christmas, Kittler, and Petrou [4]. It is the effective control of these structural errors which is our main concern in Section 6. For the time being we assume that structural errors can also be modeled by a uniform distribution which occurs with probability P_ϕ. In other words, if the match allocated to the node u_k is null or if a match to a dummy node is being considered, then a uniform probability P_ϕ is assigned. As we shall demonstrate experimentally in Section 7, this uniform modeling of structural errors is somewhat naive. For instance, it fails to account for simple situations such as a structurally consistent island being surrounded by a sea of background nodes. In-fact, when we develop our constraint filtering strategy, we altogether circumvent the need for a distribution describing structural errors. In the case of graph-editing, on the other hand, the uniform probability P_ϕ can be more easily understood as being the prior associated with the node deletion hypothesis. At this point, suffice to say that we introduce the distribution of P_ϕ largely for reasons of formal expediency.

Under these dual assumptions concerning the nature of matching errors the confusion probabilities appearing under the product of (6) may be assigned according to the following distribution rule

$$P\left(f\left(u_k\right)|v_k\right) = \begin{cases} P_\phi & \text{if } f\left(u_k\right) = \phi \text{ or } v_k = \text{dummy} \\ \left(1-P_\phi\right)\left(1-P_e\right) & \text{if } f\left(u_k\right) = v_k \\ \left(1-P_\phi\right)P_e & \text{if } f\left(u_k\right) \neq v_k \end{cases} \tag{7}$$

As a natural consequence of this distribution rule the joint conditional probability is a function of three distinct quantities

- The first variable is the total number dummy nodes padding the model graph relation S_i to the same size as the data graph super-clique C_j. This quantity, which is denoted by $\psi(\Gamma_j, S_i)$, is equal to the size difference between the structure preserving mapping S_i and the data graph super-clique C_j, i.e., $\psi(\Gamma_j, S_i) = |C_j| - |S_i|$. It conveys only information concerning the structural compatibility of the super-cliques under comparison and is not directly related to the overall consistency of match.
- The second quantity is the number of null-labels assigned to the nodes of the data super-clique C_j, which we denote by $\Psi(\Gamma_j)$. This quantity conveys information concerning the degree of noise or clutter residing in a data-graph super-clique.
- The third quantity measures the consistency of match between the non-null elements of the data-graph relation Γ_j and the non-dummy nodes of the model-graph relation S_i. This quantity counts the number of locations at which there are conflicts between the current matching assignment Γ_j residing on the super-clique C_j and those non-dummy assignments demanded by the relational mapping S_i. In other words, our measure of relational consistency is the Hamming distance $H(\Gamma_j, S_i)$ between the non-null matches belonging to Γ_j and the unpadded components of the feasible relational mapping S_i. Since it excludes dummy nodes and null-matches, the Hamming distance $H(\Gamma_j, S_i)$ is directly concerned neither with the measurement of structural differences nor the clutter fraction.

With the three variables counting the number of null matches, dummy nodes and matching errors, the joint conditional probability of matching is equal to

$$P\left(\Gamma_j|S_i\right) = P_\phi^{\psi\left(\Gamma_j,S_i\right)+\Psi\left(\Gamma_j\right)} \left[\left(1-P_\phi\right)P_e\right]^{H\left(\Gamma_j,S_i\right)} \left[\left(1-P_\phi\right)\left(1-P_e\right)\right]^{\left|C_j\right|-H\left(\Gamma_j,S_i\right)-\Psi\left(\Gamma_j\right)-\psi\left(\Gamma_j,S_i\right)} \tag{8}$$

Collecting terms together, we arrive at the following expression for the conditional matching probability

$$P\left(\Gamma_j|S_i\right) = \left(\left(1-P_\phi\right)\left(1-P_e\right)\right)^{|C_j|} \left[\frac{P_\phi}{\left(1-P_\phi\right)\left(1-P_e\right)}\right]^{\psi\left(\Gamma_j,S_i\right)+\Psi\left(\Gamma_j\right)} \left[\frac{P_e}{\left(1-P_e\right)}\right]^{H\left(\Gamma_j,S_i\right)} \tag{9}$$

All that now remains is to specify our model for the joint priors of the structure preserving mappings drawn from the dictionary, i.e., $P(S_i)\ \forall S_i \in \Theta_j$. Here we assume that each of the mappings is equi-probable, i.e.,

$$P\left(S_i\right) = \frac{1}{\left|\Theta_j\right|}.$$

Drawing on this model, our final development is to re-write the power-terms appearing in (9) using the natural exponential function. Substituting the resulting simplification of (9) into (5), we arrive at the following expression for the joint matching probability

$$P\left(\Gamma_j\right) = \frac{K_{C_j}}{\left|\Theta_j\right|} \sum_{S_i \in \Theta_j} \exp\left[-\left(k_e H\left(\Gamma_j, S_i\right) + k_\phi\left\{\psi\left(\Gamma_j, S_i\right) + \Psi\left(\Gamma_j\right)\right\}\right)\right] \tag{10}$$

where

$$K_{C_j} = \left[\left(1-P_e\right)\left(1-P_\phi\right)\right]^{\left|C_j\right|}.$$

The two exponential constants appearing in the above ex-

pression are related to the matching-error probability and the null match probability. It is a straightforward matter to show that $k_e = \ln\frac{(1-P_e)}{P_e}$ and $k_\phi = \frac{(1-P_e)(1-P_\phi)}{P_\phi}$. The probability distribution may be regarded as providing a natural way of softening the hard relational constraints operating in the model graph. The most striking and critical feature of the expression for $P(\Gamma_j)$ is that the consistency of match is gauged by a series of exponentials that are compounded over the dictionary of consistently mapped relations. As we shall demonstrate in the next section, it is this feature that distinguishes it from alternatives reported in the literature [3], [4]. Each relational mapping contributes a single exponential to the probability of match. In consequence, our method is able to operate in a robust manner when the space of relational mappings is large. We illustrate this point experimentally in Section 7 where we compare our exponential consistency measure with one based upon a linear sum of Hamming distance. Moreover, the importance of the different constraints is graded by a natural symbolic measure of relational affinity, namely Hamming distance; relational mappings of large Hamming distance contribute insignificantly while those of small Hamming distance dominate.

Finally, we construct a global relational consistency measure Q so as to model the joint prior $P(f)$. Our philosophy is to approximate the joint prior by averaging the super-clique matching probabilities over the nodes of the data graph. The consistency of the current match is therefore measured by the quantity

$$Q = \frac{1}{|V_1|}\sum_{j\in V_1} P(\Gamma_j) \tag{11}$$

With this measure to hand we are now in a position to make iterative updates in the matching function f using the MAP decision rule given in (4). However, before we proceed to detail algorithms for controlling and updating the match, we pause to consider how the relational consistency measure Q relates to some alternatives reported elsewhere in the literature.

5 Approximating the Consistency Functional

The novel feature of our relational consistency measure Q is its compound exponential structure; this distinguishes it from many alternatives described in the literature. Our aim in this Section is to elucidate the relationship with these alternative approaches by considering suitable limiting conditions for the exponentials appearing in (10).

5.1 The Hard Limit

We commence our comparative investigation by considering the conditions under which our exponential consistency measure performs the function of counting consistently matched super-cliques. Our motivation for exploring this limit is that clique counting lies at the heart of Shapiro and Haralick's relational distance metric [26]. The Shapiro and Haralick scheme minimizes the numbers of inconsistently matched N-tuples from the graphs under match. It is therefore equivalent to our method which attempts to maximize the number of consistent super-cliques. The clique counting limit occurs when $k_e \to \infty$, i.e., when $P_e \to 0$. Under these conditions the exponential Hamming distance functions appearing in equation (10) approach their Dirac delta-function limits, i.e.,

$$P(\Gamma_j) \simeq \frac{K_{C_j}}{|\Theta_j|}\exp\left[-k_\phi\Psi(\Gamma_j)\right]\sum_{S_i\in\Theta_j}\delta\left(H(\Gamma_j,S_i)\right)\exp\left[-k_\phi\psi(\Gamma_j,S_i)\right] \tag{12}$$

The combined exponential function

$$\exp\left[-k_\phi\left(\psi(\Gamma_j,S_i)+\psi(\Gamma_j)\right)\right]$$

effectively weights those consistently matched fragments of the super-cliques according to how many dummy nodes are added or how many null-labels are assigned. Provided that $P_\phi < \frac{1}{2}$, then the larger the number of assigned null-labels or dummy nodes the smaller the weight. Shapiro and Haralick [26] allowed null insertions with zero associated penalty. In order to nullify the effect of the exponential weighting, we must choose the value of P_ϕ so that $k_\phi = 0$. This is the case provided $P_\phi = \frac{1}{2}$.

Under these conditions each super-clique contributes to the overall consistency of match Q in a binary way; inconsistent super-cliques make a zero contribution while consistent super-cliques contribute an amount

$$\frac{K_{C_j}}{|\Theta_j|}.$$

In other words, the probability distribution $P(\Gamma_j)$ counts the number of zero Hamming distance super-cliques, and the goal of relational matching process is to maximize the quantity

$$Q_H = \frac{1}{|V_1|}\sum_{j\in V_1}\frac{K_{C_j}}{|\Theta_j|}\sum_{S_i\in\Theta_j}\delta\left(H(\Gamma_j,S_i)\right) \tag{13}$$

As a result of performing updates to maximize Q_H, the consistently matched portion of the graph will consist of those super-cliques for which there is a relational mapping of zero Hamming distance; the set of nodes satisfying this condition is denoted by $\Omega = \{i \in V_1 \mid \exists S_i \in \Theta_j . st . H(\Gamma_j, S_i) = 0\}$. Moreover, most of the computations involved may be effected by table look-up to identify the existence or otherwise of zero Hamming distance configurations.

It is immediately obvious that the resulting consistency measure will lead to deadlocks in the update process. If the consistency of individual super-cliques can not be restored by single label replacements, then the update process will be unable to iteratively improve the overall quality of match by deterministic means. The only way to escape such deadlocks is via an expensive stochastic optimization scheme [12]. It is also clear that the method can not be applied effectively unless there are a substantial number of consistent super-cliques present in the initial match used to

seed the relaxation scheme. Moreover, since the consistent super-cliques are unique in their label structure, the iterative overheads associated with the matching process are considerably reduced. Once a super-clique has zero Hamming distance, there is no further improvement to be gained by label updates; the nodes contained within it may be removed from the set under consideration for further updating. One way of controlling the matching process is to rank super-cliques according to their Hamming distance. In this way islands formed by abutting zero Hamming distance super-cliques may be identified. The available computational resources may be concentrated at the perimeters of the islands to extend the regions of consistency. In this way the updating process may be propagated as a "brush fire" from the consistently matched portions of the graph, obviating the need for exhaustive iterative search. In-fact, this type of optimization process is closely related to tabu-search [14], [15]; a detailed report describing the use of tabu search in conjunction with the framework outlined in this paper can be found in the recent account of Williams, Wilson and Hancock [30].

5.2 The Soft Limit

The second interesting limit of our graph-matching criterion is the case when the exponentials may be approximated in a linear way. This is the case when $k_e H(\Gamma_j, S_i) \rightarrow 0$. This limit occurs when the constant k_e approaches zero, i.e., when the error probability $P_e \simeq \frac{1}{2}$. Under these conditions the compounding of exponentials under the summation in (10) can be approximated by a linear sum of Hamming distances over the set of relational mappings,

$$P(\Gamma_j) \simeq \frac{K_{C_j}}{|\Theta_j|} \exp\left[-k_\phi \Psi(\Gamma_j)\right] \times \sum_{S_i \in \Theta_j} \exp\left[-k_\phi \psi(\Gamma_j, S_i)\right]\left[1 - k_e H(\Gamma_j, S_i) + \ldots\right] \tag{14}$$

Again it is interesting to consider the case in which there is zero penalty associated with null matches and dummy nodes, i.e., $P_\phi = \frac{1}{2}$. Under these conditions, maximization of the global consistency measure Q is achieved by minimizing the following aggregate Hamming-distance measure between the two graphs

$$Q_A = \sum_{j \in V_1} \frac{K_{C_j}}{|\Theta_j|} \sum_{S_i \in \Theta_j} H(\Gamma_j, S_i) \tag{15}$$

Several features of this approximation deserve further comment. In the first instance, the minimization of a linear function of Hamming distance as a measure of total error is precisely the function performed by the Hopfield memory in the associative recovery of patterns. In the case of a binary label set, the computation of compound Hamming distance can be effected on simple linear neurons. This observation immediately raises issues relating to the limitations of the method. In the Hopfield memory, it is well known that the linear compounding of Hamming distance restricts the number of patterns that can be effectively recalled due to inter-pattern competition [11]. The linear approximation may therefore only be anticipated to work effectively when the set of relational mappings is small.

An obvious way of overcoming this storage difficulty is to prune all but the salient structure preserving mappings from the dictionary. One heuristic yet effective strategy is to delete structure preserving mappings which correspond to matches for which there is little observational evidence. In fact, the weighting of Hamming distance according to the exponential size difference factor $\exp[-k_\phi \psi(\Gamma_j, S_i)]$ provides a natural way of excluding structurally inconsistent matches. An alternative is to tailor a specific pruned dictionary to each individual node using information concerning the initial matching probabilities.

6 Controlling the Matching Process

In the previous section we concentrated on how consistency is measured. On its own, attempting to optimize the global consistency measure Q, will only have the effect of rectifying initialization errors. Of greater practical importance is the identification and correction of structural errors caused by the addition of noise or clutter. Our aim in this section is to describe three contrasting strategies which exploit the consistency concept to overcome structural errors during the matching process [31], [32]. The first of these draws directly upon the rather naive uniform probability model described in Section 4 and attempts to label clutter segments with a null label ϕ while simultaneously correcting initialization errors. The second strategy uses the classical concept of the association graph [2]. Operating without regard to null matches, i.e., with $P_\phi = 0$, we first attempt to correct initialization errors and then construct the association graph for the optimal set of matches. Clutter is then identified in a post-processing operation that involves searching for maximum cliques of the association graph [31]. The final strategy is to use our consistency measure to identify nodes that are potential structural errors. By deleting these nodes from the graph and reconfiguring the edge-set we perform an active matching process [32]. This latter technique aims to simultaneously optimize not only the labeled match of the graph but also to restore its intrinsic relational structure so as to rectify both sources of error described in Section 4.

6.1 Null Labeling by Optimization

Our aim here is to simultaneously both rectify initialization error and label extraneous nodes in an optimization process that aims to maximize the a posteriori probability of match $P(f \mid \mathcal{A}_1, \mathcal{A}_2)$. Details of the development of the MAP update scheme have already been furnished in Section 3. Substituting for $P(\Gamma_j)$ from (10) into the MAP update rule of (4), the assignment of matches is realized on a node-by-node basis as follows:

$$f(u) = \arg\max_{v \in V_2} \frac{P(u,v|\underline{x}_u^1, \underline{x}_v^2)}{P(u,v)} \sum_{j \in C_u} \frac{K_{C_j}}{|\Theta_j|} \sum_{S_i \in \Theta_j} \exp\left[-\left(k_e H(\Gamma_j, S_i) + k_\phi \{\psi(\Gamma_j, S_i) + \Psi(\Gamma_j)\}\right)\right] \quad (16)$$

In applying the decision rule of (16), the control of the parameters P_e and P_ϕ becomes an issue of critical importance in regulating the number of null matches. Analysis of (10), reveals that it does not become energetically favorable for an erroneous match to switch to the null category until

$$P_e < \frac{P_\phi}{1-P_\phi}.$$

Since our ultimate aim is to eliminate initialization error, we take the view that the parameter P_e should be treated in the spirit of a control variable much like the temperature in an annealing process [12]. Accordingly, we reduce P_e to a small terminal value according to some deterministic iteration dependent schedule. The probability of structural errors P_ϕ, on the other hand, is a fixed property of the graphs and should remain at a constant value that reflects prior expectations concerning the level of structural corruption.

As we will demonstrate experimentally in Section 8, the null-labeling idea is most effective when the data represents a subgraph of the model. Here there is little structural corruption of the structure preserving mappings in the dictionary and the probability distribution appearing in (8) is relatively justifiable. In other words, the process can accommodate initialization errors but not structural errors.

6.2 Constraint Filtering Applied to the Association Graph

The optimization scheme described in the previous subsection iteratively corrects initialization errors while identifying extraneous nodes as belonging to the null category. One of the difficulties with this scheme lies in controlling the value of the null match probability P_ϕ. If this is set too high then there is a danger of labeling the entirety of nodes as null. If, on the other hand, the value of P_ϕ is too low then a substantial fraction of extraneous nodes will remain mismatched. An alternative strategy which is less demanding in terms of parameter control is to locate a set optimal matches with disregard to the null label ϕ. Clutter is then removed by constraint filtering aimed at identifying consistently matched subgraphs. One strategy that suggests itself is to draw on the association graph idea of Barrow and Burstall [2]. Nodes in the association graph represent putative matches; the edges indicate that the matched nodes are connected in both the data and model graphs. Consistently matched subgraphs are maximal cliques of the association graph.

Our strategy in exploiting the association graph to identify structural errors is as follows. We first locate a set of matches so as to optimize the configuration probability given in (10) by setting the value of P_ϕ to $\frac{1}{2}$, i.e., by disregarding the null-label ϕ. In other words

$$f(u) = \arg\max_{v \in V_2} \frac{P(u,v|\underline{x}_u^1, \underline{x}_v^2)}{P(u,v)} \sum_{j \in C_u} \frac{K_{C_j}}{|\Theta_j|} \sum_{S_i \in \Theta_j} \exp\left[-k_e H(\Gamma_j, S_i)\right] \quad (17)$$

The net anticipated effect of this process is to eliminate initialization errors. However, residual inconsistencies persist in the form of the erroneous matching of segments for which no feasible match exists. We therefore require a post-processing step which discards unmatchable elements to a null category. This is applied once the optimization process reaches convergence and no further iterative improvements in the consistency of match occur. In keeping with the philosophy of the association graph we commence by forming a new graph $G_1' = (V_1', E_1')$ which contains the consistently labeled portions of G_1. To form G_1' we first eliminate arcs whose mapping does not appear in G_2 to construct the new edge-set $E_1' = \{(u_1,u_2)|(f(u_1), f(u_2)) \in E_2\}$. We next remove disjoint nodes which are no longer connected by an arc; these nodes have no support and may therefore be considered to be unmatchable. The resulting node set is denoted by $V_1' = \{u|(u,v) \in E_1', v \in V_1\}$. Disjoint nodes are consigned to the null category, i.e., $u \notin V_1' \Rightarrow f(u) = \phi$. The graph G_1' now consists of a number of internally connected yet potentially disjoint patches, in which all the nodes satisfy the arc consistency constraint. Suppose that Λ_i denotes the index-set of one of these disjoint graph partitions. If there are p such partitions, then

$$V_1' = \bigcup_{i=1}^{p} \Lambda_i$$

and there are no interconnecting edges, i.e., if $i \neq j$ then

$$(\Lambda_i \times \Lambda_j) \cap E_1' = \varnothing.$$

If correct matching were the only process which generated consistency, then G_1' would contain only the correctly labeled portion of the graph. However, a small amount of spurious consistency is generated from local matches between unmatchable segments and regions of G_2. Since the probability of accidentally forming a partition of size $|\Lambda_i|$ in the initial labeling is equal to $|V_2|^{-|\Lambda_i|}$ we gauge the overall consistency of each partition by the number of nodes contained within it. In practical matching applications the regions of spurious consistency are small, typically much smaller than correctly labeled regions, and in this case it is sufficient to reject those regions whose size falls below some threshold value T

$$|\Lambda_i| < T \Rightarrow \forall u \in \Lambda_i [f(u) = \phi] \quad (18)$$

In this way we can effectively filter-out relational units which are not consistently matched. This constraint filtering process can be regarded as a rudimentary form of graph editing that is applied to locate consistent subgraphs once the optimization of consistent matches has converged. Good results can therefore be anticipated if the matching process is concerned with locating fragmented subgraphs.

If, on the other hand, the structural errors uniformly permeate the graphs then a finer editing strategy is required.

6.3 Active Matching With Graph-Edit Operations

Under conditions of extreme structural error, relational fragmentation will severely limit the matching methods described in Sections 6.1 and 6.2. The basic problem originates from topological disruption of the consistent relations residing in the dictionary by the addition of structural errors. For instance, if the relational structure being matched is a Delaunay graph representing a Voronoi tessellation of the image plane, then a single additional clutter node may create many additional graph edges [32]. One way of overcoming this problem is to edit-out extraneous nodes. In this way the relational structure of the data graph is reconstructed in an active matching process. This process is illustrated in Fig. 2. As originally proposed by Sanfeliu and Fu [24], graph editing was realized as a cost minimization process. There were separate edit operations and associated costs for node deletion or insertion, edge deletion or insertion, and, for edge or node relabelings. Because we confine ourselves to Delaunay graphs, our edit operations are considerably simplified. The reason for this is that the edge-sets associated with node insertions or deletions are automatically generated from the Voronoi tessellation. In other words, our edit process only involves node insertion or deletion, and, node relabeling.

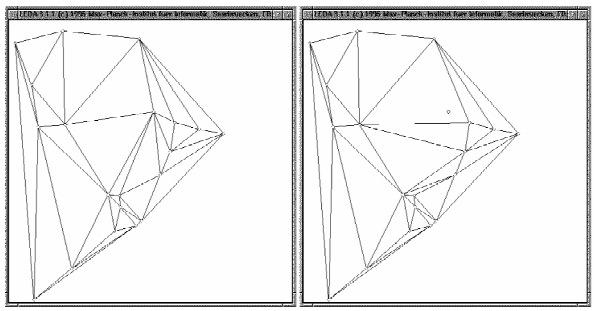

Fig. 2. An example of Delaunay graph editing by node deletion. The original graph is shown on the left, while the edited graph resulting from a node deletion is shown on the right.

We have recently shown how this process of active graph reconfiguration or editing can be realized as a MAP estimation process similar to that described in Section 3 [32]. However, rather than confining itself purely to node relabeling, the algorithm now encompasses the possibility of node insertions or deletions together with the implied modification of the Delaunay edge-set. Details of the derivation are outside the scope of this paper. The basic idea is to gauge the net effect of deleting a node by examining those contributions to the consistency measure that arise from modification of the super-cliques containing the node in question. This set is constructed by identifying those nodes that form a super-clique with node u in graph G_1, i.e., $C_u - \{u\}$, and determining the new super-clique set for these nodes in the reconfigured graph G_1'. We let χ_u^+ denote the super-clique set of object u in graph G_1 and χ_u^- denote the corresponding super-clique set in the reconfigured graph G_1'. With this notation the change in the MAP criterion caused by the deletion of the node u is proportional to

$$\Delta_u^- = P_\phi \sum_{j \in \chi_u^-} \frac{K_{C_j}}{\left|\Theta_j\right|} \sum_{S_i \in \Theta_j} \exp\left[-k_e H\left(\Gamma_j, S_i\right)\right] \tag{19}$$

By contrast, when considering the change in the MAP criterion caused by re-insertion of the node u it is the super-clique set χ_u^+ to which we turn our attention. The corresponding change to the MAP criterion is proportional to

$$\Delta_u^+ = \frac{P\left(u, f(u) \mid \underline{x}_u^1, \underline{x}_{f(u)}^2\right)}{P\left(u, f(u)\right)} \sum_{j \in \chi_u^+} \frac{K_{C_j}}{\left|\Theta_j\right|} \sum_{S_i \in \Theta_j} \exp\left[-k_e H\left(\Gamma_j, S_i\right)\right] \tag{20}$$

The decision criteria for node deletion or re-insertion are as follows. We delete node u provided $\Delta_u^+ < \Delta_u^-$ and reinstate it provided $\Delta_u^+ > \Delta_u^-$. In other words, the rule for assigning nodes to the null category is

$$\Delta_u^+ < \Delta_u^- \Rightarrow f(u) = \phi \tag{21}$$

Once the graph has been reconfigured or edited in this purely structural manner, labeling consistency must be restored to those portions of the graphs to have undergone structural modification. This is effected by modifying the labels assigned to the nodes of the data-graph. Adhering to our MAP philosophy, the labeling consistency is maintained provided that the mapping function f is updated in the following way

$$f(u) = \arg\max_{v \in V_2} \frac{P\left(u, v \mid \underline{x}_u^1, \underline{x}_v^2\right)}{P\left(u, v\right)} \sum_{j \in \chi_u^+} \frac{K_{C_j}}{\left|\Theta_j\right|} \sum_{S_i \in \Theta_j} \exp\left[-k_e H\left(\Gamma_j, S_i\right)\right] \tag{22}$$

The graph reconfiguration process of (21) and the node relabeling operation of (22) combine to perform a function analogous to Sanfeliu and Fu's graph edit operations [24]. Rather than requiring the specification of a series of heuristic cost terms, the dual operations of node deletion or insertion, and, node relabeling are both regulated by the null match probability P_ϕ and the error probability P_e. Since the parameter P_e fulfills the role of a control variable which imposes incremental hardening on the constraints residing in the dictionary, our structural error model has only one free parameter P_ϕ. This provides certain advantages in terms of ease of control over the heuristic edit costs of Sanfeliu and Fu [24].

7 Sensitivity Analysis

There are two aspects to our comparative sensitivity study. In the first instance, we are interested in comparing the effectiveness of our compound exponential consistency measure with its approximate counterparts. In this way we offer some comparison with the methodology adopted by of Shapiro and Haralick [27], and, the plethora of Hopfield-like implementations of relational matching [28], [29]. The

second aspect of our experimental study concerns the different strategies adopted in the control of structural errors. Here we give some direct performance assessment of the three alternative strategies described in Section 6.

In order to provide some objective performance measures, we have conducted the bulk of our experiments using synthetic relational graphs. In order to embark on this study, we have seeded Voronoi tessellations from random dot patterns and computed the associated Delaunay graph. The nodes of the graph are therefore the random dots, while the arcs indicate that the associated Voronoi regions are adjacent to one-another. In order to simulate the effects of clutter and segmental dropout, we have both added dots at random locations and deleted random dots from the patterns used to seed the Voronoi tessellation. This has the effect of corrupting the topology of the associated Delaunay graph.

In addition to the structural errors introduced by random dot insertions and deletions, we have also added random initialization errors. Associated with each node in the Delaunay graph is a uniformly distributed random unary attribute vector. This unary attribute is used to compute initial match probabilities. Uncertainties in the measurements used in the discrete relaxation scheme are modeled by randomly perturbing the unary attributes of the data-graph with a Gaussian error distribution so as to produce a specified fraction of initial matching errors. The initial matching probabilities are computed from exponential distributions of the Mahalanobis distance between attribute-vector pairs computed using an estimate of the variance-covariance matrix Σ, i.e.:

$$P\left(u, v \mid \underline{x}_u^1, \underline{x}_v^2\right) = \left(1 - P_\phi\right) \frac{\exp\left[-\frac{1}{2}\left(\underline{x}_u^1 - \underline{x}_v^2\right)^T \Sigma^{-1}\left(\underline{x}_u^1 - \underline{x}_v^2\right)\right]}{\sum_{w \in V_2} \exp\left[-\frac{1}{2}\left(\underline{x}_u^1 - \underline{x}_w^2\right)^T \Sigma^{-1}\left(\underline{x}_u^1 - \underline{x}_w^2\right)\right]} \tag{23}$$

Where appropriate, initial null matches are assigned a probability $P(u, \phi) = P_\phi$ which is independent of the attribute values. Our assumed model for the matching priors is one of uniformity over the model graph i.e., $P(u, v) = (|V_2|)^{-1}$. In our simulation studies, we use only a single scalar attribute at each node. For the majority of the experiments described here, the unary attribute errors are generated so that the level of initialization error from the winner-take-all assignment of matches is approximately 50 percent.

7.1 Consistency Measure

Our first set of experiments are aimed at comparing the effectiveness of the different limiting approximations to the compound exponential consistency measure developed in Section 5. The two cases of interest are the hard-limit and the linear limit; the respective limits under which the approximations apply are, $P_e \to 0$ and $P_e \to \frac{1}{2}$.

7.1.1 The Hard-Limit

Turning our attention first to the case of the hard limit, Fig. 3 shows the final fraction of correct matches as a function of the fraction of added clutter as defined above. The dotted curve is the result of applying the approximation of (13), while the solid curve is the result of applying the exponential consistency measure. For each of the points in the graph, the matching algorithm has been initialized in a state where 50 percent of the matches are incorrectly assigned. The exponential form of the consistency criterion consistently outperforms its hard-limiting counterpart until the fraction of graph corruption exceeds 60 percent. Moreover, even when the graphs are perfect realizations of one-another, the hard limit is only capable of rectifying some 10 percent of the initialization errors. In other words, the hard limit is not only sensitive to structural errors in the graphs, it is also extremely limited in its capacity to recover from significant levels of initialization error.

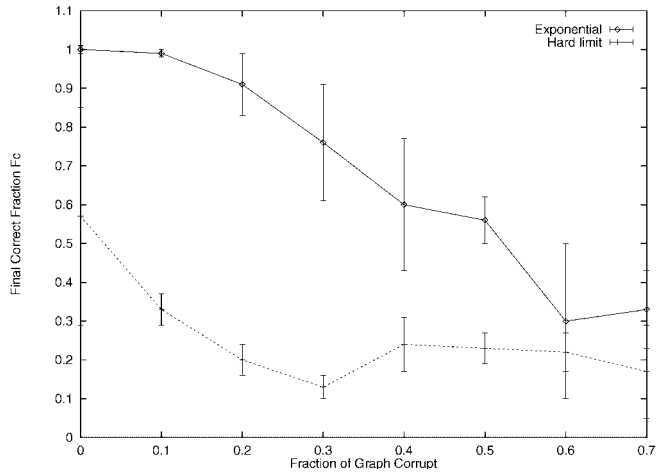

Fig. 3. Comparison of performance of the hard-limit (dotted curve) with the exponential criterion (solid curve).

The sensitivity of the hard limit to initialization error is investigated more thoroughly in Fig. 4. This plot shows the final matching accuracy as a function of the fraction of initially correct matches for structurally consistent graphs. The points on the scatter-gram represent the results of applying the matching process using the hard-limit consistency measure, while the dotted-curve shows the result obtained using the exponential consistency measure. The exponential version of the consistency measure is capable of recovering a fully consistent match even when as few as 15 percent of the initial matches are correct. By contrast, the hard limit is only effective when the number of initialization errors is less than 50 percent. Moreover, the drop-off in performance is very rapid. In fact, the performance of the hard-limit is crucially dependent on the presence of a significant number of correctly matched super-cliques from which to seed the iterative matching process.

7.1.2 The Linear Limit

To commence our investigation of the linear form of the consistency measure we have performed some experiments on uncorrupted graphs. Here the only source of error is due to mis-initialization. Fig. 5 shows the final fractional accuracy as a function of graph size when the initialization error is 50 percent. The dotted curve is the result obtained using the exponential consistency measure while the lower dashed-curve is the result of applying the linear approximation. Whereas the exponential form recovers a completely consistent labeling irrespective of graph size, the linear form results in a degradation in performance which increases with graph size. In the latter case, the match is effectively a random configuration of symbols.

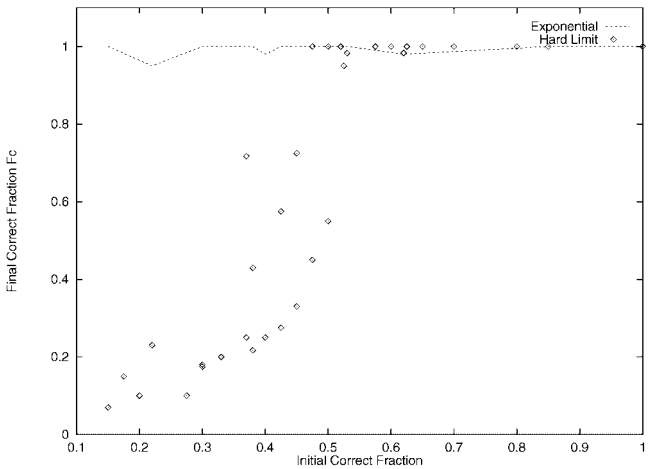

Fig. 4. Recovery from initialization errors for the hard-limit (diamond points) and the exponential criterion (dotted curve).

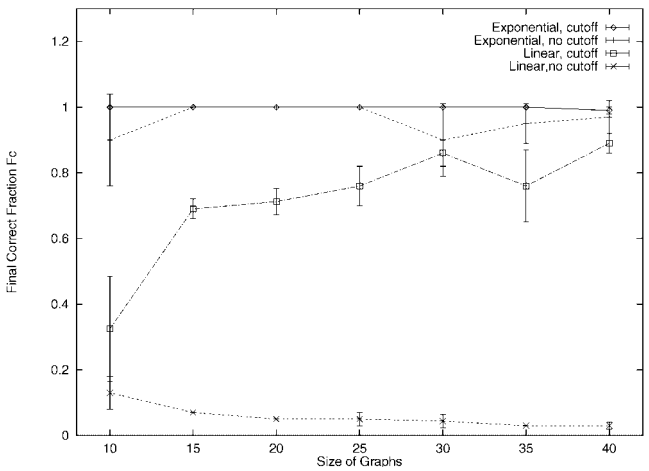

Fig. 5. Comparison of linear and exponential forms: exponential (dotted curve); linear (dashed curve); pruned exponential (solid curve); pruned linear (dot-dashed curve).

By applying the dictionary pruning strategy, the results described above can be dramatically improved. Each dictionary is pruned by excluding 80 percent of the mappings on the basis of initial match probabilities. In the case of the linear consistency measure which is shown as a dot-dashed curve, final matching accuracies of between 60 percent and 90 percent can be obtained. Moreover, as the graph size increases the relational mappings become less ambiguous and the matching performance increases. As shown by the solid curve, in the case of the exponential consistency measure, the pruning process results in a small improvement in performance. This effect is attributable to the fact that the pruning strategy reduces matching ambiguities for large graphs.

Having established that the linear consistency measure can be rendered operable under conditions in which errors are dominated by mis-initialization, we proceed to investigate the effects of structural corruption. Fig. 6 compares the pruned linear consistency measure with its exponential counterpart using the data described in Section 7.1.1. The graphs used in this study are relatively large (40 nodes) and hence represents the domain where the linear approximation can be anticipated to be effective in controlling initialization errors. The dotted-curve in Fig. 6 reveals that although inferior to the exponential consistency measure, the pruned linear form considerably outperforms the hard limit (Fig. 3). This is an interesting observation. It indicates that even drawing on Hamming distance in a suboptimal way is more effective than the relatively coarse consistent clique counting of Shapiro and Haralick [27].

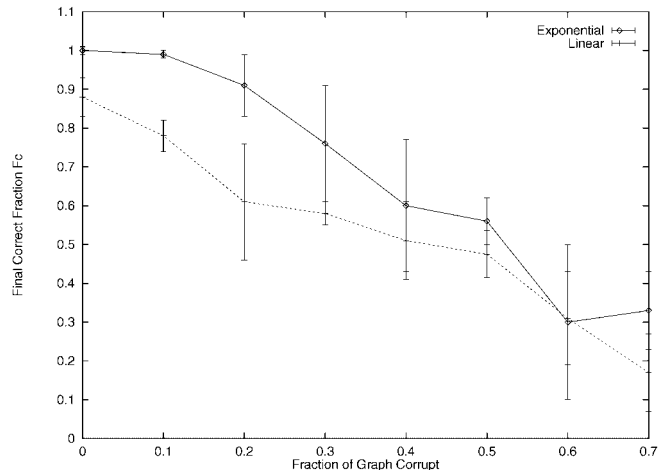

Fig. 6. Comparison of pruned linear and exponential consistency measures.

7.2 Clutter Control

Our aim here is to offer a quantitative comparison of between different strategies for controlling structural errors. Accordingly, we have performed a series of simulation experiments to establish the effective operating limits of the three strategies for rectifying structural errors described in Section 6. Fig. 7 shows the fraction of the original graph correctly matched as a function of the fraction of added clutter. The solid line represents the result of active matching with graph-edit operations (see Section 6.3). The dashed line is the result of applying constraint filtering to the association graph of the optimal matches (see Section 6.2). Finally, the dotted line is the result of null-labeling (see Section 6.1). It is clear that the active matching method consistently outperforms both null-labeling and constraint filtering. Errors do not appear until the fractional corruption

exceeds 20 percent. Even when the clutter fraction is as high as 50 percent, then 80 percent matching accuracy is achievable. By contrast, the two remaining methods are 10-20 percent more susceptible to error. However, it is interesting to note that the constraint filtering technique has a tangible performance edge over null-labeling. This is largely attributable to the difficulties encountered in controlling the parameter of the null match. For comparison, Fig. 8 shows the fraction of residual noise after matching; this represents the fraction of added noise which is not assigned to the null category. Again, the best performance is obtained using graph-edit operations, with null-labeling delivering the poorest performance.

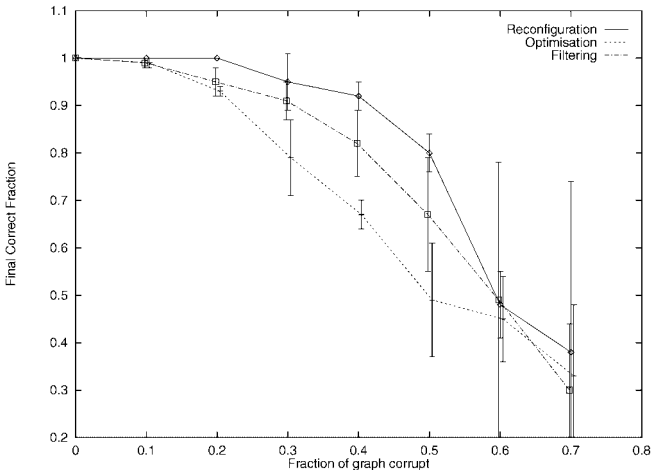

Fig. 7. Fraction of graph correctly matched.

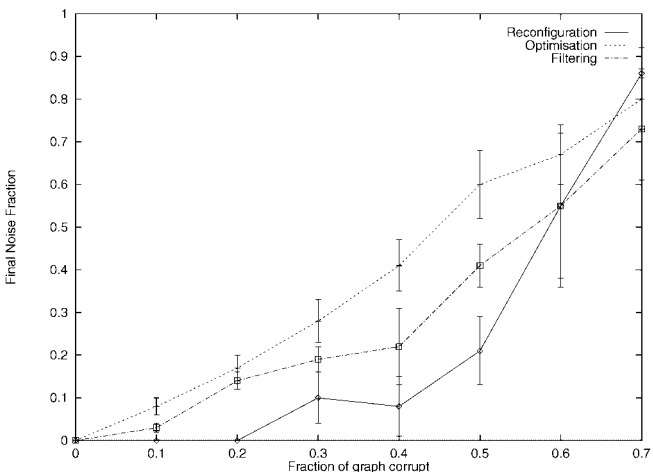

Fig. 8. Residual noise fraction.

8 Real-World Imagery

Having presented a sensitivity analysis for the various realizations of our matching scheme, in this section we will confront real-world data. We provide two examples of the application of our matching framework. In both applications we are concerned with matching aerial image data against digital map data. The first example the application involves the matching of road networks detected in infrared line scan images. Here the fidelity of the ground truth information is good and, additionally, the segmentation process is effective in providing a stable relational description for matching. As a result there are few structural errors and in consequence the control of clutter is not an issue of critical importance. The second application is far more demanding. It involves the matching of hedge structures detected in synthetic aperture radar images. Here there are two factors which make the effective control of structural errors an issue of critical importance. Firstly, the imagery is dominated by speckle noise which results in significant clutter. The second limitation stems from the poor quality of the ground truth map information for hedge structures. This is due largely to the fact that they are the artifacts of vegetation and are subject to the vagaries of changing land-use.

8.1 Aerial Infrared Images

We are interested in comparing the effectiveness of the various consistency measures described in Section 5 when applied to the matching aerial scenes against digital map data. For the experimental aspects of this study we take the application of matching road networks detected in the image data against their corresponding map representation. The image data under study was obtained using an infrared line scan sensor. In addition to the transformational differences due to changes of viewpoint and scale, there is barrel distortion in the horizontal direction due to the geometry of the line-scan process. This effect is most marked at the image periphery. Image scan in the vertical direction is controlled by aircraft motion across the ground. Irregularities in the aircraft flight-path due to banking or turbulence introduce vertical scanning distortions. An example image is shown on the left-hand side of Fig. 9 while the right-hand image is the digital map data available for matching. The left-hand image in Figs. 10 and 11 represents digital map data while the right-hand image shows the road network segmented from the image data using a relaxational ridge detector [19].

Our matching of the two scenes is based on finding correspondences between the T-junctions which delineate the road network. According to our graph-based abstraction of the matching process the nodes represent T-junctions while the arcs signify the existence of a connecting line-structure. Initial match probabilities are assigned using (22) on the basis of angle differences between the line-structures forming the junctions. Figs. 10 and 11 illustrate the matching results obtained using the compound exponential criterion of (10) with the MAP update rule of (16); correct matches are shown in Figure 10 while matching errors are contained in Fig. 11. Initially, of the 158 T-junctions in the map 35 match correctly, 66 match in error and 57 are null matched. After application of the iterative update procedure described in Section 6.1, 72 match correctly, 77 match to the null category and only 9 are incorrectly matched; the fraction of correct matches has therefore increased from 0.22 to 0.46. An illustration of the iterative improvement resulting from the application of the discrete relaxation scheme specified by (16) are illustrated in Table 1. Each row lists the results obtained with successive iterations of the scheme with the label error probability reduced according to a deterministic schedule. The columns in Table 1 show

the numbers T-junctions which are correctly matched, incorrectly matched or null matched. Also listed in Table 1 is the prevailing value of the label-error probability.

TABLE 1
RESULTS OF MATCHING

Iteration	Correct	Incorrect	Null	P_e
0	46	63	0	-
1	57	52	0	0.300
2	68	41	0	0.234
3	67	42	0	0.182
4	73	36	0	0.142
5	75	34	0	0.110
6	76	33	0	0.086
7	76	25	8	0.067
8	79	17	13	0.052
9	76	16	17	0.041
10	80	11	18	0.032

Under similar initial matching conditions, when the hard limiting criterion of (13) is applied the corresponding fraction of correct matches in the final labeling is 0.34. Although this result is inferior, the match is still useful. By contrast the soft limiting criterion of (15) is completely ineffective in locating any kind of useful global match; it is dominated by null matches and matching errors.

(a) (b)

Fig. 9. Aerial infrared image (left) and digital map (right)}

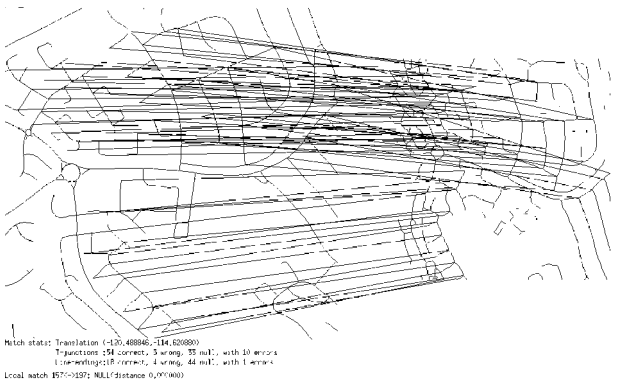

Fig. 10. Correct matches between the map (left) and the infrared data (right).

8.2 Synthetic Aperture Radar Data

Evaluation of our methodology for rectifying structural errors is concerned with matching hedge structures segmented from synthetic aperture radar (SAR) images against their cartographic representation in a digital Ordnance Survey map. The hedge structures responsible for radar reflections appear as intensity ridges in the raw data. As with the infrared data, we again perform feature detection by applying a relaxation operator to the output of a set of orientational line-detectors. The primitives used in matching are linear segments extracted from the raw feature contours. Fig. 12 illustrates the processing operations applied to the raw SAR image together with the results of applying the graph edit operations described in Section 6.3. Fig. 12a. shows the original SAR image. The extracted hedge features used in matching are shown in Fig. 12b. The relational graphs used in the matching process are Delaunay triangulations computed from the linear segments in the SAR data and the digital map. Figs. 12c and 12d show the model and data graphs used for matching.

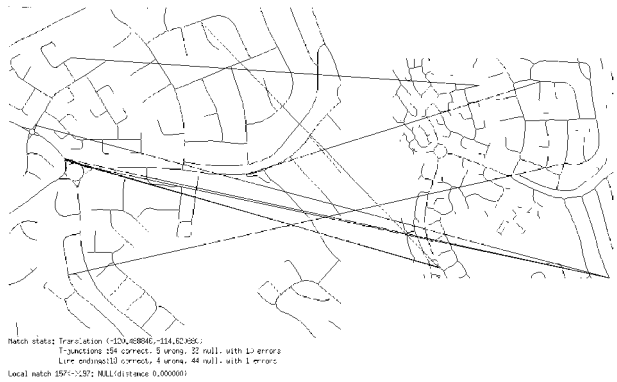

Fig. 11. Incorrect matches between the map (left) and the infrared data (right).

The initial matches between the linear segments extracted from the SAR data and their map representation are established on the basis of the affinity between the vectors of the angle differences between model and data lines using (22).

The experimental matching study is based on 95 linear segments in the SAR data and 30 segments contained in the map. However only 23 of the SAR segments have feasible matches within the map representation. Fig. 12e shows the initial matches for the line segments in the SAR data; the black lines are correct matches while the gray lines are matching errors. With the same coding scheme Fig. 12f shows the final result once the iterative graph-edit process has converged. Here those lines edited from the graphs have been deleted from the figure. Comparing Figs. 12e and 12f, it is clear that the main effect of the graph-edit operations, has been to delete the majority of the clutter segments from the SAR data graph. To give some idea of relative performance merit, in the case of the initial matching configuration 20 of the 23 matchable segments are correctly identified with 75 incorrect matches, while after application of the graph-editing method the final graph contains 19 correct matches, only 17 residual clutter with 59 nodes deleted from the data graph.

The comparison of the three strategies for controlling structural errors on the SAR data is summarized in Table 2. The table lists the fraction of nodes correctly matched F_c

and together with the fraction of mis-labeled error-nodes F_n. A perfect result would be one for which $F_c = 1$ and $F_n = 0$. Of the three schemes, the null-labeling approach gives the poorest results. Distinguishing between the graph-edit technique and the constraint filtering technique is less clear-cut. Although the constraint filtering method results in a larger fraction of correctly matched nodes, there is also a greater fraction of residual structural errors. One way of reconciling these figures is to note that the matchable line segments in Fig. 12b, represent a number of fragmented subgraphs. It is under these conditions that the constraint filtering technique is the most applicable. As demonstrated by the simulation studies in Section 7, the active graph-editing technique offers the best performance when structural errors permeate genuine structure in a uniform way.

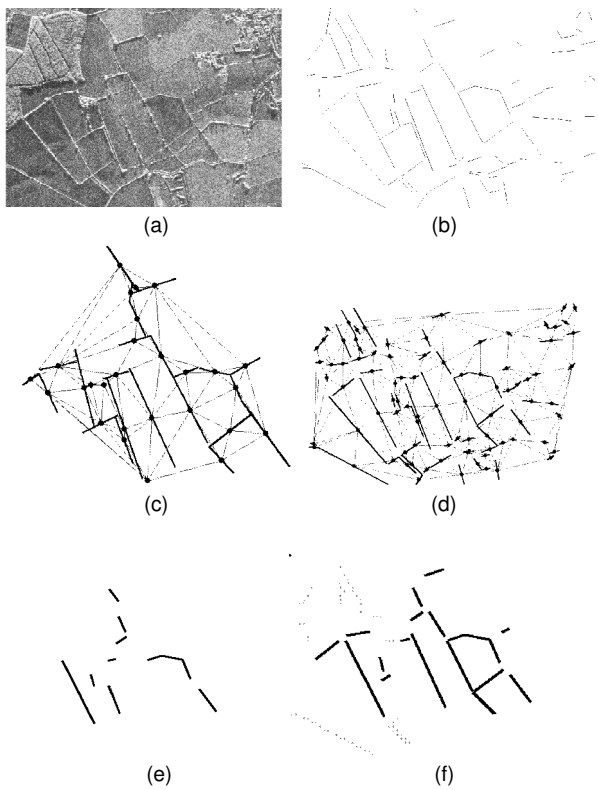

Fig. 12. Graph editing. (a) Original image. (b) Extracted line segments. (c) Model graph. (d) Data graph. (e) Initial match. (f) Final edited match.

TABLE 2
SUMMARY OF MATCHING RESULTS FOR THE SAR DATA

Control strategy	F_c	F_n
Constraint filtering	0.96	0.56
Null labeling	0.57	0.77
Graph edits	0.77	0.47

9 Conclusions

Our main contributions in this paper have been twofold. Firstly, we have provided a uniform Bayesian framework for understanding a diverse family of relational matching algorithms. The second aspect has been to provide some extensive experimental comparisons. From the theoretical perspective, we have explored a number a ways in which relational consistency can be gauged. Based on an experimental study, the most effective of these appears to be using a sum of exponential Hamming distance functions. The linear and delta-function limits of the exponentials are both suboptimal in their performance. Having established the most effective choice of consistency measure, we have provided some comparative study of different strategies for accommodating structural inexactness. These include null labeling, graph editing and the use of the association graph. Here we found that the best performance was delivered by a graph-editing technique. However, when combined with the compound exponential consistency measure in a hybrid matching algorithm, the association graph also produced encouraging results.

References

[1] H.G. Barrow and R.J. Popplestone, "Relational Descriptions in Picture Processing," *Machine Intelligence*, vol. 6, pp. 377–396, 1971.
[2] H.G. Barrow and R.M Burstall, "Subgraph Isomorphism, Matching Relational Structures and Maximal Cliques," *Information Processing Letters*, vol. 4, pp. 83–84, 1976.
[3] K. Boyer and A. Kak, "Structural Stereopsis for 3D Vision," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 10, pp. 144-166, 1988.
[4] W. Christmas, J. Kittler, and M. Petrou, "Probabilistic Feature-Labeling Schemes—Modeling Compatibility Coefficient Distribution," *Image and Vision Computing*, vol. 14, pp. 617–625, 1996.
[5] W. Christmas, J. Kittler, and M. Petrou, "Structural Matching in Computer Vision Using Probabilistic Relaxation," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 17, pp. 749–764, 1995.
[6] S. Dickinson, A. Pentland, and A. Rosenfeld, "3D Shape Recovery Using Distributed Aspect Matching," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 14, pp. 174–198, 1992.
[7] O.D. Faugeras and K.E. Price, "Semantic Labeling of Aerial Images Using Stochastic Relaxation," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 3, pp. 633–642, 1981.
[8] A.M. Finch, R.C. Wilson, and E.R. Hancock, "Matching Deluanay Graphs," *Pattern Recognition*, vol. 30, pp. 123–140, 1997.
[9] D.B. Fogel, "An Introduction to Simulated Evolutionary Optimization," *IEEE Transactions on Neural Networks*, vol. 5, pp. 3–14, 1994.
[10] P.J Flynn and A.K. Jain, "CAD-Based Vision—From CAD Models to Relational Graphs," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 13, pp. 114-132, 1991.
[11] E. Gardner, "Structure and Metastable States in the Hopfield Model," *J. Physics, vol.* A19, pp. L1,047–L1,052, 1986.
[12] D. Geman and S. Geman, "Stochastic Relaxation, Gibbs Distributions and Bayesian Restoration of Images," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 6, pp. 721-741, 1984.
[13] D. Geiger and F. Girosi, "Parallel and Deterministic Algorithms From MRF's: Surface Reconstruction," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 13, pp. 401–412, 1991.
[14] F. Glover, "Genetic Algorithms and Tabu Search—Hybrids for Optimization," *Discrete Applied Mathematics*, vol. 49, pp. 111-134, 1995.
[15] F. Glover, "Tabu Search for Nonlinear and Parametric Optimisation (With Links to Genetic Algorithms)," *Discrete Applied Mathematics*, vol. 49, pp. 231–255, 1995.
[16] S. Gold and A. Rangarajan, "A Graduated Assignment Algorithm for Graph Matching," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 18, pp. 377–388, 1996.
[17] E.R. Hancock and J. Kittler, "Discrete Relaxation," *Pattern Recognition*, vol. 23, pp. 711–733, 1990.
[18] E.R. Hancock and J. Kittler, "A Bayesian Interpretation for the Hopfield Network," *1993 IEEE Int'l Conf. Neural Networks*, pp. 341–346, 1993.
[19] E.R. Hancock, "Resolving Edge-Line Ambiguities Using Probabilistic Relaxation," *Proc. IEEE Computer Society Conf. Computer Vision and Pattern Recognition*, pp. 300–306, 1993.

[20] R. Horaud and T. Skordas, "Stereo Correspondence Through Feature Grouping and Maximal Cliques," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 11, pp. 1,168–1,180, 1989.
[21] W. Kim and A.C. Kak, "3-D Object Recognition Using Bi-Partite Graph Matching Embedded in Discrete Relaxation," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 13, pp. 224–251, 1991.
[22] B.T. Messmer and H. Bunke, "Efficient Error-Tolerant Subgraph Isomorphism Detection," D. Dori and A. Bruckstein, eds., *Shape, Structure and Pattern Recognition*, pp. 231-240, 1994.
[23] A. Rosenfeld, R. Hummel, and S. Zucker, "Scene Labeling by Relaxation Operations," *IEEE SMC*, vol. 6, pp. 420–433, 1976.
[24] A. Sanfeliu and K.S. Fu, "A Distance Measure Between Attributed Relational Graph," *IEEE SMC*, vol. 13, pp. 353–362, 1983.
[25] K. Sengupta and K.L. Boyer, "Organizing Large Structural Model-Bases," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 17, pp. 321–332, 1995.
[26] L. Shapiro and R.M. Haralick, "Structural Description and Inexact Matching," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 3, pp. 504–519, 1981.
[27] L. Shapiro and R.M. Haralick, "A Metric for Comparing Relational Descriptions," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 7, pp. 90-94, 1985.
[28] P. Simic, "Constrained Nets for Graph Matching and Other Quadratic Assignment Problems," *Neural Computation*, vol. 3, pp. 268-291, 1991.
[29] P.N. Suganathan, E.K. Teoh, and D.P. Mital, "Pattern Recognition by Graph Matching Using Potts MFT Networks," *Pattern Recognition*, vol. 28, pp. 997–1,009, 1995.
[30] M.L. Williams, R.C. Wilson, and E.R. Hancock, "Deterministic Search Strategies for Relational Graph Matching," *Advances in Energy Minimisation Methods for Computer Vision*, Lecture Notes in Computer Science, to appear.
[31] R.C. Wilson, A.N. Evans, and E.R Hancock, "Relational Matching by Discrete Relaxation," *Image and Vision Computing*, vol. 13, pp. 411–422, 1994.
[32] R.C. Wilson and E.R. Hancock, "Relational Matching With Dynamic Graph Structures" *Proc. Fifth Int'l Conf. Computer Vision*, pp. 450–456, 1995.
[33] R.C. Wilson and E.R Hancock, "A Bayesian Compatibility Model for Graph Matching," *Pattern Recognition Letters*, vol. 17, pp. 263–276, 1996.
[34] A.K.C. Wong and M. You, "Entropy and Distance of Random Graphs With Application to Structural Pattern Recognition," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 7, pp. 509–609, 1985.
[35] D. Yang and J. Kittler, "MFT Based Discrete Relaxation for Matching High Order Relational Structures," *Proc. 12th Int'l Conf. Pattern Recognition*, vol. 2, pp. 219-223, 1994.

Richard C. Wilson is currently a research associate in the Department of Computer Science at the University of York. Dr. Wilson was awarded an open scholarship to read physics at St. John's College, University of Oxford, graduating with first-class honors in 1992. Between 1992 and 1995, he undertook research at the University of York on the topic of relational graph matching, for which he was awarded the DPhil degree. He has published some 35 papers in journals, edited books, and refereed conferences. His research interests are in high-level vision, scene understanding, volumetric image analysis, and structural pattern recognition.

Edwin R. Hancock gained his BSc in physics in 1977 and PhD in high-energy nuclear physics in 1981, both from the University of Durham, England. After a period of postdoctoral research working on charm-photo-production experiments at the Stanford Linear Accelerator Centre, he moved into the fields of computer vision and pattern recognition in 1985. Between 1981 and 1991, he held posts at the Rutherford-Appleton Laboratory, the Open University, and the University of Surrey. Since 1991, Dr. Hancock has been a lecturer in computer science at the University of York, where he now leads a group of 15 researchers in the areas of computer vision and pattern recognition. He has published about 125 refereed papers in the fields of high-energy nuclear physics, computer vision, image processing, and pattern recognition. He was awarded the 1990 Pattern Recognition Society Medal. Dr. Hancock serves as an associate editor of the journal *Pattern Recognition* and has been a guest editor for the *Image and Vision Computing Journal*. He is currently guest-editing a special edition of the *Pattern Recognition* journal devoted to energy minimization methods in computer vision, and pattern recognition. He chaired the 1994 British Machine Vision Conference and has been a program committee member for several national and international conferences.